Integrating Legacy Tools and Data Sources

Jon Dukes-Schlossberg
Staff Scientist
Lockheed Martin Missiles & Space
Advanced Software Technology
Sunnyvale, California
408-742-4796
jon.dukes-schlossberg@lmco.com

Eamon Barrett
Consulting Scientist
Lockheed Martin Missiles & Space
Advanced Software Technology
Sunnyvale, California
408-743-2901
eamon.barrett@lmco.com

Abstract

Under DARPA and internal funding, Lockheed Martin has been researching information needs profiling to manage information dissemination as applied to logistics, image analysis and exploitation, and battlefield information management. We have demonstrated key software tools to aid in the dissemination of data based on a user's needs and the changing world environment. The Lockheed Martin effort has included assisting in the design and development of user information needs models, software that instantiates those information needs models with world information, and software that assists in the creation and management of those profiles. We have applied this profiling technology to several applications. Under DARPA's Joint Logistics Program, we are disseminating key transportation logistics data to disparate users. Specifically, we are providing a unitary login to combat support, transportation, and theater, Defense Logistics Agency, and supplier inventories. Under DARPA's Battlefield Awareness and Data Dissemination program, we are improving both the accuracy and timeliness of "smart information push" to the warfighter. By applying these techniques, "smart push" delivers just the right data, to just the right place, at just the right time. Under internal Lockheed Martin funding, we are investigating the use of information profiles to facilitate user interactions with an image analysis and exploitation system. This system integrates several disparate legacy applications communicating over a commercial publish-subscribe middleware package. We are using Lockheed Martin's commercial Intelligent Library System (ILS) to provide image archiving and retrieval. This paper concentrates on the image analysis and exploitation system..

Keywords

User profiling, information dissemination, information integration

1. Introduction

Information is available from many different computer sources and in many different formats: unstructured and semistructured text (e.g., news wires, intelligence reports, etc.), structured data in database systems (e.g., logistics and inventory databases), and image and video data (e.g., weather maps from satellites). Navigating this space to obtain relevant information in a timely manner would improve the efficacy and efficiency of problem solving/decision making. The information is out there, but finding, combining, and receiving updates to the information in a timely manner are the problems.

To address this, Lockheed Martin researchers have been investigating information integration and dissemination technology for many years. Beginning with foundational efforts defining agent communication languages, agent architectures, ontologies, and infrastructure tools, we have evolved this work into focusing on user information needs modeling and how to achieve "just the right information to just the right user at just the right time." This focusing has entailed defining information needs profiling concepts: architectures, representation languages, creation, and continuous updating techniques.

Form SF298 Citation Data

Report Date *("DD MON YYYY")* 00001999	**Report Type** N/A	**Dates Covered (from... to)** *("DD MON YYYY")*

Title and Subtitle Integrating Legacy Tools and Data Sources	**Contract or Grant Number**
	Program Element Number
Authors	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Lockheed Martin Missiles & Space Advanced Software Technology Sunnyvale, California	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 6	

Accurately capturing the information needs of a user and pushing that information to a user is of great interest on the World Wide Web where we are "drowning in information but starving for knowledge." Companies such as Marimba, PointCast, Talarian, and TIBCO to name just a few are all working to enable this vision. The current state-of-the-practice involves user's subscribing to "information channels." Any information published to that channel would be forwarded to the subscribing user. We seek to vastly expand this concept by constructing dynamic information needs profiles that are instantiated at run-time as opposed to compile time. These profiles capture information needs at a much finer level of detail than channel subscriptions. Further, profiles are re-instantiated over time.

DARPA's Intelligent Integration of Information (I3) program, underway since 1992, seeks to facilitate end-user and application program access to information by extracting, integrating, and abstracting information from data sources. The goals of an I3-based system are to provide seamless access to heterogeneous information sources and allow legacy application programs to operate in a cooperative, coordinated manner. The vision is to facilitate the interoperation of distributed intelligent agents and their access to data. We seek to be complementary to the I3 vision. Much of our work is based on agent concepts emerging originally from the I3 program.

2. Architecture

Figure 1 details our notional architecture for an information needs profiling environment. As detailed in subsequent sections, our various profiling implementations have demonstrated some of these components working together. We have not yet fully demonstrated and deployed the full architecture.

This architecture provides effortless yet accurate and reliable information needs profiling. Unlike conventional static and manual profiling, we automatically assess the information needs of users, manage them in reaction to dynamically changing situations, and assure their prompt delivery by profile optimization and end-to-end Quality of Service (QoS) estimation and negotiation with the underlying network structure. We relieve users from deciding what information they need and whether or not the information would be delivered in a timely manner. It accommodates both conditional and event-driven information needs through a flexible and expressive profile representation. The system considers the underlying QoS-related constraints in assessing information needs.

Visualized as a collection of distributed and collaborating software agents our architecture, as shown in Figure 1, integrates innovative yet scaleable multi-disciplinary technologies. Information needs are assessed based on the user's mission, identity, and current world status, and composed into an information profile by the Information Needs Profiler. The composed information profile can be further customized through the Profile Editor. The Event Manager manages events associated with profiles, providing the ability to dynamically adapt the information needs to a changing environment. Multiple profiles are merged and optimized by the Profile Processor based on priorities of information and channel and bandwidth availability. The QoS Manager estimates the required QoS based on the up-to-date status of the underlying network and priorities of information. The QoS Manager also provides feedback to the Information Needs Profiler to assist in selecting alternative choices of information. Our design includes building on a commercial publish-susbcribe middleware package such as Active Software's ActiveWorks or Talarian Corporation's SmartSockets .

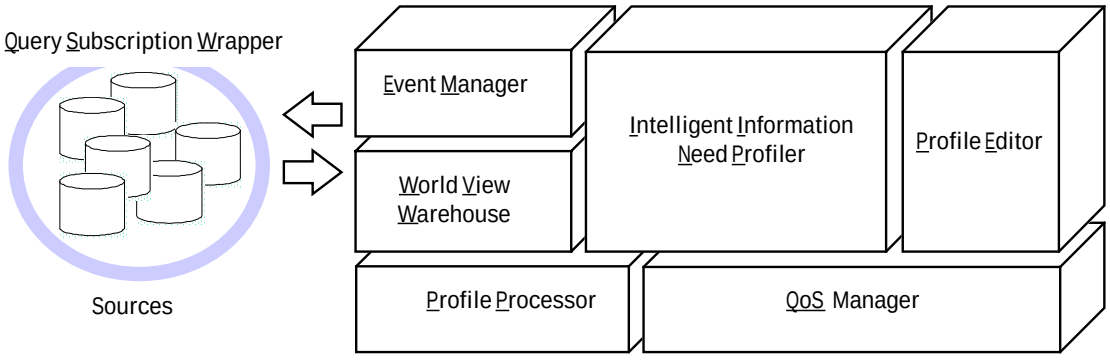

Figure 1. Notional information needs profiling architecture. Our view of information needs profiling integrates to produce accurate profiles with minimal user input and maintenance.

One goal of our profiling system is to have a user new to a task or environment boot up their machine and quickly start receiving information; a profile is constructed for them based on stored templates and the user's identity, role, location, current operation, etc. We can collect most of this information from networks and information sources; in some application domains, some user input may be required at the outset. In the military domain, we believe no user input would be required for the system to a create a profile automatically and have that profile used by the system. The user can customize this profile at their leisure.

3. Implementation

We have implemented portions of our visionary user profiling architecture in several domains. This paper describes the implementation for the Information Fusion Architecture (IFA) internal Lockheed Martin research program. Other implementations remain under development at this time. Specifically the Joint Logistics program is implementing a user profiling agent with interesting persistent query properties.

3.1 Information Fusion Architecture

The Information Fusion Architecture effort seeks to develop a framework to integrate disparate legacy systems and then develop and deploy a resulting system. The application domain for this work is image exploitation and analysis. The system under development integrates two image exploitation systems, a mapping system, the Intelligent Library System(ILS) , all communicating over the ActiveWorks commercial publish-subscribe middleware system. With the exception of ActiveWorks, the other systems are all Lockheed Martin-developed. The goals of this integration effort are:

- Demonstrate enhanced decision making and value-added through the integration of disparate tools, not originally designed to work together, but now able to cooperate and collaborate on problem solving
- Demonstrate the ease and rapidity of the integration using a standards-based distributed object approach

Image Exploitation is a fertile application domain for this work. Presently, Intelligence Analysts (IAs) are involved in a very labor intensive, high skill analysis of data with little automation or decision aid support. The enormous volumes of data can not be processed effectively by these analysts. There is growing concern that critical problems or crises may not be detected in a timely manner (e.g., the surprise to the US of the Indian nuclear test). New information and new techniques are opening imagery exploitation to many new problem sets. The potential to identify activities, events, processes, materials, etc. is exploding beyond our capacity to process and disseminate the information.

We are integrating our legacy systems with new algorithms and tools. We are developing both classified and unclassified scenarios. Figure 2 presents the architecture for our information fusion work. This diagram shows two legacy image exploitation and analysis systems, the FOCUS change detection system and the PNN (Probabilistic Neural Network) spectral processing system, integrated into a cooperative problem solving environment. The core of this system is the user profiling agent in which the information needs of users are modeled and then used to facilitate a "push" model of information dissemination; this is indicated as the "User Layer" in the Figure. With the profiling agent, the user specifies their geographic region of interest, the image processing systems to execute with images matching the geographic region of interest, and the notification method (e.g., e-mail, pager) to use when a matched image has successfully passed through the execution steps. The BADGER (Bay Area Digital GeoResource) mapping system provides a graphical tool for geographic area selection as well as providing a context for results displayed back to the user. The “Services Layer” provides the workflow processing of matched images through the available image processing tools. Other services have been implemented on other projects at Lockheed Martin (e.g., the Simulation-Based Design project has implemented notification, object discovery, mediation, and more sophisticated workflow) and we intend to integrate these capabilities into our system. The “Middleware Transport Layer” provides the messaging infrastructure support. We are using the ActiveWorks commercial publish-subscribe middleware system to provide this functionality. We have also worked with various implementations of the CORBA standard but have found the ActiveWorks product, as a commercial offering, provides the scaleability, security, and reliability we need (and documentation, as well). Finally, we use the Intelligent Library System(ILS) , a commercial image archiving and retrieval system from Lockheed Martin, to match images, store images, and assist in user notification.

3.2 Profiling Implementation

The advanced user profile agent software enables the construction and modification of a user's persistent set of interests (i.e., queries). A priori user interests and key attributes about their role, mission, status, etc. form their "profile." This user profile is then used as a basis to automatically construct and then maintain the user's information needs. This agent has the capability to define, maintain, and dynamically modify user profiles for multiple users or for a single user who is responsible for monitoring multiple sites and/or regions of interest. These pre-defined user profiles can then be disseminated over broadcast satellite and other means to the intelligence data processing nodes and, when coupled with distributed object-oriented services and databases to provide a "smart-push" tailored reporting capability to mission planning centers or directly to operational units/users in the battlefield. This agent utilizes an intelligent software agent (the Interest Agent) which, in cooperation with the ActiveWorks Event Manager, provides control of data processing and information dissemination events.

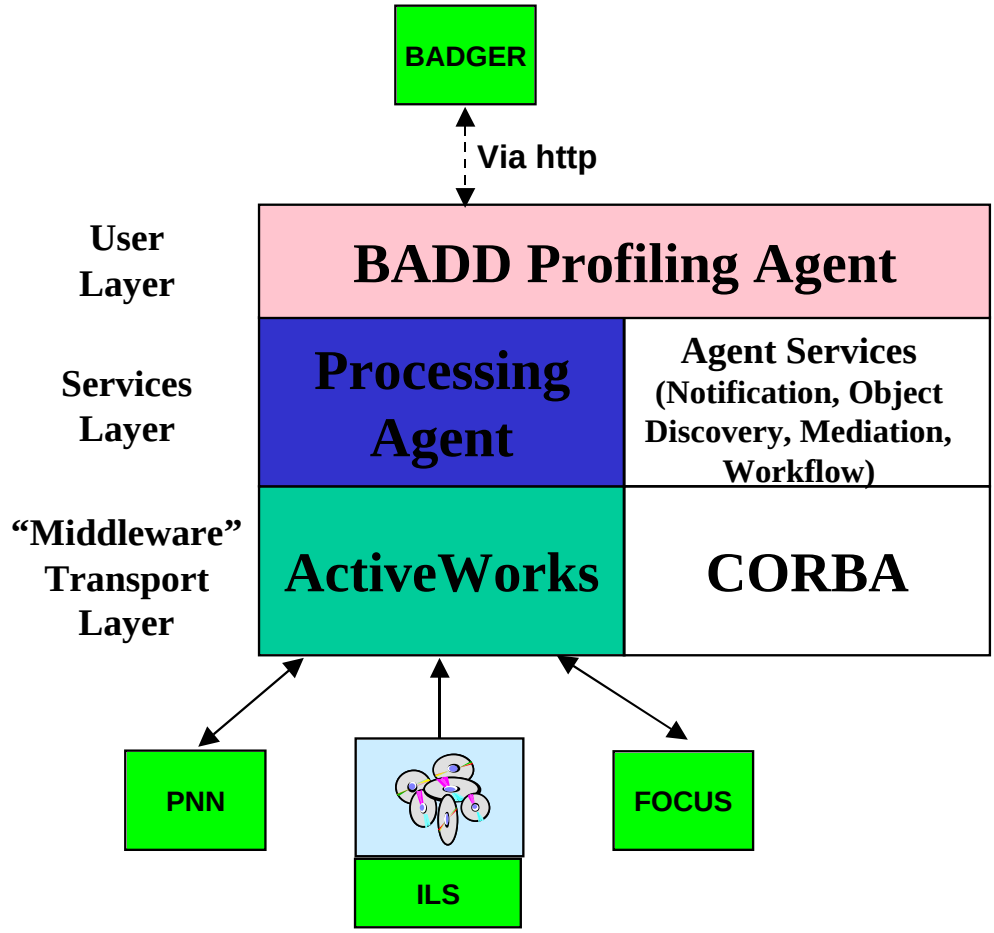

Figure 2. Information Fusion Architecture (IFA) project architecture.

Operationally, a user's information needs translates into a set of "interests." In our image exploitation domain, an interest represents a standing query for an image type over a certain area of interest (e.g., San Francisco airport), a workflow that invokes the image exploitation tools integrated into the system in possibly interesting ways, and a notification request that specifies how the user wants to be notified when the system has a result to present.

This prototype system is tightly integrated with the Intelligent Library System(ILS) . This system is a commercial image archiving and retrieval system developed by Lockheed Martin in Sunnyvale. We are exploiting its built-in image ingest, profile matching, image archiving, querying, and notification. This system has been and continues to be sold to international, government, and commercial entities.

The scenario beings with the end user interacting with the profiling agent to develop his set of information profiles. The user graphically specifies their information and this is registered in the Intelligent Library System's database. This is equivalent to specifying a persistent query for this image. The Intelligent Library System monitors an image stream

and looks for geographic (and metadata) matches between the images and its registered profiles. If a match is found, the profile agent is notified and starts the workflow associated with that profile. This workflow can be a complex specification of a set of processing steps to take. For example, a scenario we support is for a workflow to first call the FOCUS change detection system. If this system returns a certain confidence level in its belief that there's a change in this image (relative to its site model), then call a hyperspectral processing tool. If this tool detects a certain chemical signature in the image, then integrate the results and notify the user.

Notice that we're not notifying the user every time an image arrives over San Francisco airport but rather after a set of image processing has taken place on the image. Thus profiles in this context have the advantage of enabling us to focus the user's attention on images we have a belief are more relevant.

We are expanding and upgrading this infrastructure architecture to add more image processing tools and to provide more abstraction from details of the infrastructure. This last feature is key: we don't want to rewrite all the wrappers if the middleware transport system is changed to CORBA or if we add a new image archive.

We have also investigated other key attributes of profiling agents that derive from our visionary architecture presented earlier. We have done some work in automated profile creation. Given we know some information about a user, we can associate him with a class of other users and predict some of his information needs. We have applied learning techniques to this problem with some success. We have also looked at automatically monitoring a user's key attributes as well as the world environment in order to update his information needs. An example of this would be reinstantiating some of the user's standing queries.

4. Intelligent Library System (ILS)

Lockheed Martin's Intelligent Library System (ILS) offers users a full end-to-end imagery and data management solution. Workflow management, storage, retrieval, and processing throughput are all tailored to the user's environment and data types. Some key features of ILS:

- Provides significant storage capacity--users can archive and manage 5,000 terabytes or more of information in various combinations of online, nearline and offline storage
- Comes in various configurations supporting single workstations with 10-gigabyte storage to dispersed enterprise systems with petabyte capacity
- Can handle input from and output to virtually any data type to support users in a wide range of industries; can disseminate information in either electronic, or hard copy formats
- Is built on a legacy of proven performance for the U.S. Government, ILS offers its customers the same robust, continuous availability we`ve already provided to the world`s most demanding end user
- Provides a number processing and exploitation tools to help quickly identify the most meaningful information within the stored data files
- Provides a web-based user interface that can be easily customized to suit the needs of individual users; uses an open data architecture to allow users to custom tailor their systems without expensive and costly system software upgrades

The Intelligent Library System provides a data archiving and retrieval capability for the Information Fusion Architecture effort.

5. Summary

Lockheed Martin Advanced Software Technology has a broad initiative in information integration and dissemination. This technology has emerged from eight years of information integration research and addresses real problems expressed by government customers. Portions of this IFA system are to be delivered and deployed to a national customer in summer, 1999. We are developing a generic technology as demonstrated by the fact that it is now a key component on several disparate projects: logistics, battlefield information management, tool integration, collaborative engineering. We will be deploying a key user profiling testbed to a national customer in the summer, 1999.

Author Bibliography

Jon Dukes-Schlossberg is a staff scientist in the Lockheed Martin Advanced Software Technology department, Lockheed Martin Missiles & Space. Jon leads a group working on information integration and dissemination concepts for a range of military applications. He has been a Lockheed Martin employee since 1987. He holds a BSE in Computer Science and Engineering from the University of Pennsylvania and an MS in Computer Science from the University of Florida.

Dr. Eamon B. Barrett is an imagery scientist and mathematician. He has authored more than 50 technical publications in mathematics, physics, and imagery science. Eamon has over 25 years of experience conducting and directing R&D projects in imagery science, most recently at the Lockheed Martin Palo Alto Research Laboratory and the Advanced Technology Center where he is employed as Research Scientist and Computer Systems Analyst. In 1968 Eamon completed his Phd in Mathematics at Stanford University, where he specialized in differential geometry and applied mathematics.